$\boxed{5.2}$

- $\overbrace{\text{counts } (x) \text{ and proportions } (\hat{p})}^{\text{Binomials}}$ are discrete r.v.'s
- sometimes approx. with normal distr. (continuous)
- stats often use continuous r.v.'s

Ex: sample means ($\bar{x}$)
Percentiles
Proportions (%)
Std. deviations (s)

	parameter	statistic
	Pop	Sample
Mean	μ	$\bar{x}$
Prop	p	$\hat{p}$
Std. dev	σ	s
	usually unknown	found from sample

Sample means = $\overline{X}$ = avg of a sample of data

- from an SRS of size n.

  simple random sample

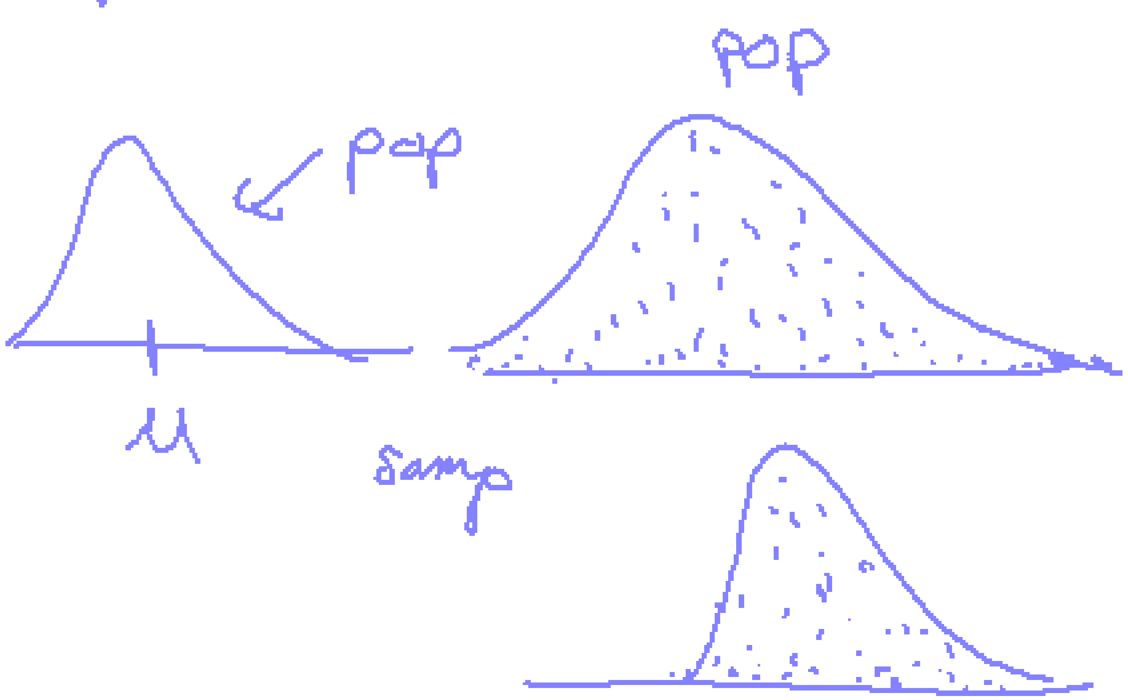
- from a population with mean μ and std. dev σ.

Distribution of X ← pop.

Shape: **Normal**

Center: μ

Spread: σ $\quad n=1$



Distribution of $\bar{X}$ ← sample

normal

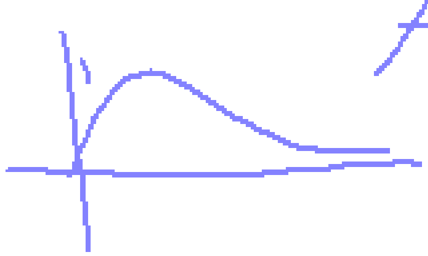
μ

$\sigma/\sqrt{n}$

$n\uparrow$ = better

$n=30$

CLT



As $n \uparrow$ from any (normal or non-normal) population, the sampling distrib. becomes more normal with

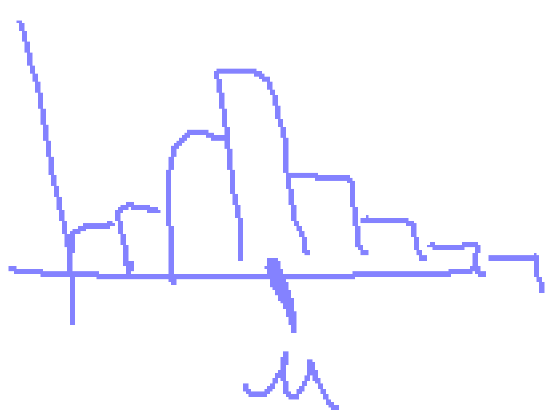
$\bar{x} \sim N(\mu, \sigma/\sqrt{n})$

histogram of $\bar{x}$'s

Ex: binomial

$n=4$

$n=20$



Check:

$n \cdot p$

$n(1-p) \geq 10$

Big enough n

- dependends on pop. shape
- more non-normal the pop, the larger n needed
- General rule:

  <u>check:</u> normal pop

  or

$B(n,p)$ μ, σ $n \geq 30$

Ex: normal ✓

$\mu = 298$

$\sigma = 3$

pop ① $P(X < 295) = \text{normcdf}(-E99, 295, 298, 3) =$ 0.1587

*samples are less variable than individual observations.

* bigger n are better!

② n=6

$P(\bar{X} < 295) = \text{normcdf}(-E99, 295, 298, \frac{3}{\sqrt{6}}) = 0.00715$

③ $\mu = 1000$ $\sigma = 20.5$

ⓐ $P(X < 975) = \text{normalcdf}(-E99, 975, 1000, 20.5)$

pop.
$n = 1$

$= 0.11132$

ⓑ $n = 45$

check
$n = 45 \geq 30$

$P(\bar{X} < 975) = \text{normalcdf}\left(-E99, 975, 1000, \frac{20.5}{\sqrt{45}}\right)$

≈ 0

μ
σ

normal ✓

① a) $P(x<2) = \text{normalcdf}(-E99, 2, 2.17, 0.11) = 0.0611$

$n \geq 30$

$n=10$

b) $P(\bar{x}<2) = \text{normcdf}(\text{"}, \text{"}, \text{"}, \frac{0.11}{\sqrt{10}}) = 5.122 \times 10^{-7}$

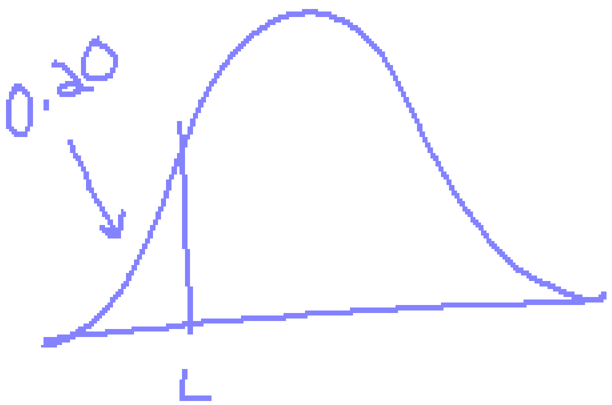
normal ✓

② a) $P(x>68{,}500) = 0.0327$

b) $P(\bar{x}>68{,}500) = 1.147 \times 10^{-4}$

$n=4$

invnorm(probability below, mean, std. dev.)



$P(\bar{x} < L) = 0.20$

---

* Distribution = shape
center
spread

$N(\mu, \sigma)$